Combined Artificial Bee Colony Algorithm and Machine Learning Techniques for Prediction Online Consumer Repurchase Intention

Anil Kumar[a*], Manoj Dash[b], Gaurav Kabra[c], Eswara Krishna Mussada[d]
Prashant Singh Rana[e]

[a]School of Management
BML Munjal University, Gurgaon, Haryana, India
Email ID: anilror@gmail.com

*Corresponding Author

[b]Indian Institute of Information Technology and Management, Gwalior
Gwalior, Madhya Pradesh, India
Email ID: manojctp@gmail.com

[c] Department of Management Studies
Xavier Institute of Management, Bhubaneswar, India
Email ID: kabraddm@iitr.ac.in

[d] School of Engineering and Technology
BML Munjal University, Gurgaon (Haryana), India
Email: eswar_mussada@yahoo.com

[e] Computer Science and Engineering Department
Thapar University Patiala, Punjab, India
Email ID: psrana@gmail.com

Combined Artificial Bee Colony Algorithm and Machine Learning Techniques for Prediction Online Consumer Repurchase Intention

Abstract

A novel paradigm in the service sector i.e. services through the web is a progressive mechanism for rendering offerings over diverse environments. Internet provides huge opportunities for companies to provide personalized online services to their customers. But prompt novel web services introduction may unfavorably affect the quality and user gratification. Subsequently, prediction of the consumer intention is of supreme importance in selecting the web services for an application. The aim of study is to predict online consumer repurchase intention and to achieve this objective a hybrid approach which a combination of machine learning techniques and Artificial Bee Colony (ABC) algorithm has been used. The study is divided into three phases. Initially, shopping mall and consumer characteristic's for repurchase intention has been identified through extensive literature review. Secondly, ABC has been used to determine the feature selection of consumers' characteristics and shopping malls' attributes (with > 0.1 threshold value) for the prediction model. Finally, validation using K-fold cross has been employed to measure the best classification model robustness. The classification models viz., Decision Trees (C5.0), AdaBoost, Random Forest (RF), Support Vector Machine (SVM) and Neural Network (NN), are utilized for prediction of consumer purchase intention. Performance evaluation of identified models on training-testing partitions (70-30%) of the data set, shows that AdaBoost method outperforms other classification models with sensitivity and accuracy of 0.95 and 97.58% respectively, on testing data set. This study is a revolutionary attempt that considers both, shopping mall and consumer characteristics in examine the consumer purchase intention.

Keywords: Artificial bee colony algorithm; classification; consumer; K-fold cross validation; prediction sensitivity.

1. Introduction

With the exponential growth of e-services (electronic services), business around the globe are trying to improve their competitive advantage by focusing their resources on the virtual business environment (Nissen and Sengupta, 2006). Mohanty et al. (2010) also opined that efficient business processes and automation have become the foremost priority for business firms. E-services are independent of language and platform, therefore anyone can access them from anywhere on the globe. Lu (2001) suggested several benefits of e-services from the perspective of business firms such as, increase competitive strength, channels of communication to the customers etc.

Purchasing has traditionally been at the convenience of the customer but increase in options, both online and offline, has triggered a change of perception in the minds of the customers, encouraging them to explore many options. Therefore, typically a customer might prefer an online purchase to an offline purchase if the variety of products is more, or it offers better convenience. These are some of the major contributors that trigger a shift from offline purchase intention to online purchase intension. Another major contributor which triggers this shift is the concept of utility maximization. With the boom in the e-commerce sector, both nationally and globally, customers now look for more alternatives and thus are exhibiting preference for more online participation than ever before.

The challenge for the service provider is the dynamic buying behavior patterns of consumers. It is said that when consumers adopt new technologies, their behaviors change (Danaher et al., 2003; Zinkhan and Watson, 1998). Many business firms utilize information technologies such as data mining techniques to extract customer's data in order to validate their strategy plan before implementation (Balabanovic & Shoham, 1997).

In today's competitive environment, loyalty and intention towards the service provider play a very vital role in customer satisfaction. Although it is a challenge to judge and guess online customer intention and their requirements, it can be analyzed through past data and customer intention when there is no individual interaction between sellers and buyers (Kumar and Dash, 2014). The other subfields of computer science, such as data mining techniques, have been utilized in business, helping enterprises to support knowledge discovery and decision making (Seng and Chen, 2010). Examples are: product mix analysis, market segmentation, customer segmentation, direct marketing, fraud detection, churn analysis, inventory analysis etc. However, research on online consumer repurchase intention has been limited in the area of consumer marketing and rarely seen a combination of Artificial Bee Colony and data mining techniques has been used by researchers in this context.

E-services can successfully compete with their offline counterparts if they understand the numerous influencing factors of electronic purchasing intention of the customer. There are numerous studies investigating the relation between website customer satisfaction and quality during online shopping. However, most of these studies either they investigate the relationship between shopping mall characteristics and consumer repurchase intention; or between consumer characteristics and consumer repurchase intention. Specific research which considered simultaneously both, shopping mall and consumer characteristics in order to examine the consumer repurchase intention is lacking. Therefore, this study is a revolutionary attempt to fill gap from literature. In nutshell, this study aims to predict online consumer intention on the basis of the characteristics of consumers and shopping malls. To achieve this, ABC Algorithm has been employed for the section of features. ABC is swarm-based algorithm and a tool for intelligent optimization which has enough capabilities to find good solution within a reasonable running time (Karaboga & Basturk, 2008; Karaboga & Akay, 2009). In features selection problem, ABC is good performer algorithm and gives better results as compare to other population-based algorithms (Karaboga & Akay, 2009,). For prediction of consumer purchase intention, along with ABC, several classification models based on intelligent techniques namely, Decision Trees (C5.0), AdaBoost, Random Forest (RF), Support Vector Machine (SVM) and Neural Network (NN).

The work has been organized in seven Sections. Second section discusses the conceptual framework of the research model. Third section designates the dataset, ABC algorithm, and description of machine learning models. Fourth section details the methodology adopted to achieve the objectives. Fifth section presents the model evaluation, while sixth section deliberates the results obtained. Conclusion of the study has been given in the last section.

2 Conceptual framework of the research model

There are a number of factors that affect online consumer buying behavior, in this study these factors divided into two parts 1) Consumers' characteristics and 2) Shopping malls' characteristics that are discussed widely in the following sections.

2.1 Consumers' characteristics

Existing research indicates that shopping value derived from online shopping of a consumer explains the characteristics of consumers (Chan et al., 2007; Lee et al., 2009; Kumar and Dash, 2014). Consumer characteristics play a central role in driving the dynamics of information seeking (Johnson et al., 2004) that, in turn, influences online buying behavior of the products.

Attractive features of products make positive impact on the orientation of consumers (Herbes and Ramme, 2014). Online customers like to buy products with attractive features for excitement and to try new things (Maniak et al., 2014). Product customization aims to ensure that such customers find new and interesting products according to their level and expectations (Lysonski and Durvasula, 2013).

Tremendous changes have recognized in consumption pattern, purchasing behaviour and brand consciousness of consumers from the period of post liberalization. Brand conscious customers tend to buy luxurious, classy and well established brand products that are extensively advertised (He et al., 2012). Such online shoppers know the value of branded products (Lysonski and Durvasula, 2013). They prefer well-known brands, globally advertised, expensive and best-selling. They are loyal towards particular brands and are perfectionists (Azad et al., 2014; Lysonski and Durvasula, 2013). Novelty of products can give a feel of excitement to the consumers, which gets reflected in their consumption (Blake et al., 2003; Jin, 2013). At present, due to the flow of huge information related to new brands, new products and new stores, the consumer feels psychologically burdened. New brands/products attract consumer but the consumer feels confused and is reluctant to try the products (Rizwan et al., 2013; Yang et al., 2014). Innovativeness (INN) is an impetus for the adoption of new products or services. Innovativeness plays an important role in regulating consumers' attitude towards online purchases and their future intention to purchase, directly affecting their intention (Citrin et al., 2000).

With extensively advertised product, less risk-averse consumers try to separate the good signs of poor products, even as the high risk-averse consumers watch for different viewpoints, or turn to shortcut buying approaches, comparable to, the use of price, brand and store cues (Chen et al., 2009; Gao et al., 2012). High risk-averse consumers may additionally turn to knowledge acquisition with a purpose to diminish the uncertainty related to purchases (Park and Lee, 2009; Sicilia and Ruiz, 2010). This trait is characterized by consumers who feel confused due to information overload (Gao et al., 2012; Park and Lee, 2009). Another characteristic of the online shopper is preference of brand by others (Cowart and Gold- smith, 2007; Riquelme and Rom´an, 2014). Likeness of the products by others can impact their purchase intention before making the purchase online (Iyengar et al., 2009; Riquelme and Rom´an, 2014). Factors often discussed in online shopping are empowering consumers (Ellis-Stoll et al., 1998; Pires et al., 2006). Empowerment means switching suppliers on the lookout for better value propositions (Pires et al., 2006). Empowered consumers are anticipated to engage in assessment shopping due to the fact they have got all types of understanding and so they constantly search for higher worth propositions (Pires et al., 2006).

2.2 Shopping malls' characteristics

The perceptions about an online store can significantly affect people's purchase intention (Kim and Gupta, 2009; Merlo et al., 2012) and that perception can create an image in their mind set and directly impact of their purchasing behavior (Merlo et al., 2012). Good corporate reputation strengthens customers' cross-buying intentions, enhancing the expectations of quality of service. It also leads to expectations of reduced costs, and increase in knowledge and confidence in the organization's commitment to consumers (Hung et al., 2012; Jeng, 2008, 2011; Verhagen and Van Dolen, 2009). The degree of consumer satisfaction about the reputation of shopping malls depends on their global recognition and centralized distributed reputation systems (Childers et al., 2002; Hung et al., 2012; Verhagen and Van Dolen, 2009).

In electronic purchases, consumers are more concerned about the return policy of the service provider and also about whether transactions at the online store are error-free (Field et al., 2004; Gretzel and Park, 2010). Consumers make online purchases for convenience, to save time; even 24*7 online shopping mall facilities are more advantageous for them (Gretzel and Park, 2010; Kumar and Dash, 2014). This means that consumers are concerned about the time and energy spent on shopping (Field et al., 2004).

With growing customer realization about social issues, showing their orientation towards society. Markets are assessed on the basis of degree of acceptance by the consumer of the concept of consumer citizen and consumers buy any item or service not only for personal satisfaction, but also for responsiveness to societal (and environmental) well-being (Johnston, 2001). Such consumers are socially conscious consumers who share their personal information on a public platform with the aim to bring social change through their purchasing power (Webster, 1975, Wu et al., 2013). It is speculated that there is a substantial relation between online consumer decision-making style and comparison shopping acceptability (Park, 2007). Nevertheless, consumer style characterized by perfectionism wants to evaluation-retailer to get "perfect" choices (Gretzel and Park, 2010; Kumar and Dash, 2014). This shows that they see these tools as valuable in organizing and helping alternative selection. Table 1 lists the features for online consumer buying intention prediction.

Table 1 Features for online consumer intention prediction

S. No.	Features	References
1.	Varieties of products (X_1)	Azizi and Makkizadeh (2012); Cheong and Chin (2012); Vyas and Sisodia (2013); Colantone and Crin`o (2014); Kochukalam and Peters (2016)
2.	Attractive features of products (X_2)	Rieger (2012); Li et al. (2013); Maniak et al. (2014); Herbes and Ramme (2014); Rajeev and Rekha (2016).
3.	Customize products (X_3)	Merle et al., (2010); Lee and Chang (2011); Thirumalai and Sinha (2011); Bright and Daugherty (2012); Wang and Tseng (2013); Kang and Lee (2014); Mahdjoubi et al. (2014)
4.	Brand loyalty (X_4)	He et al. (2012); Lysonski and Durvasula (2013); Azad et al. (2014); Kumar and Dash (2016)
5.	The best quality products (X_5)	Lysonski and Durvasula (2013), Azizi and Makkizadeh (2012)
6.	Stickiness of Brands (X_6)	Jin (2013); Citrin et al. (2000); Rizwan et al. (2013), Goldsmith and Hofacker (1991); Yang et al. (2014)
7.	Fun of Electronic purchasing (X_7)	Lee and Shin (2014); Kumar and Dash (2015); Kumar and Dash (2016)
8.	Confusion availability of many online store (X_8)	Azizi and Makkizadeh (2012); Gao et al., (2012)
9.	Excessive information (X_9)	Chen et al., (2009); Park and Lee (2009), Sicilia and Ruiz (2010); Gao et al. (2012)
10.	Friends influence (X_{10})	Cowart and Goldsmith (2007); Iyengar et al., (2009); Forbes and Vespoli (2013); Riquelme and Rom´an (2014)
11.	Likeness of brand by others (X_{11})	Iyengar et al. (2009); He et al. (2012)
12.	Value for money (X_{12})	Kim et al., (2014), Verhagen and Van Dolen (2009)
13.	Empowerment (X_{13})	Ellis-Stoll and Popkess-Vawter (1998); Park (2007); Pires et al., (2006); Gretzel and Park (2010)
14.	Fast purchasing facility (Y_1)	Kumar and Dash (2013)
15.	Save time (Y_2)	Park (2007); Gretzel and Park (2010)
16.	24*7 online facilities (Y_3)	Park (2007); Pires et al. (2006); Gretzel and Park (2010)
17.	Global recognition (Y_4)	Jeng (2011); Hung et al. (2012); Merlo et al. (2012)
18.	Online reputation (Y_5)	Kim and Gupta (2009); Verhagen and Van Dolen (2009); Hung et al. (2012); Merlo et al. (2012); Kumar and Dash

		(2016)
19.	Companies' involvement in society welfare (Y_6)	Johnston (2001); Wu et al. (2013)
20.	Price comparison (Y_7)	Park (2007); Gretzel and Park (2010); Zhou et al. (2016)

X represents consumer characteristics and Y represents shopping mall characteristics

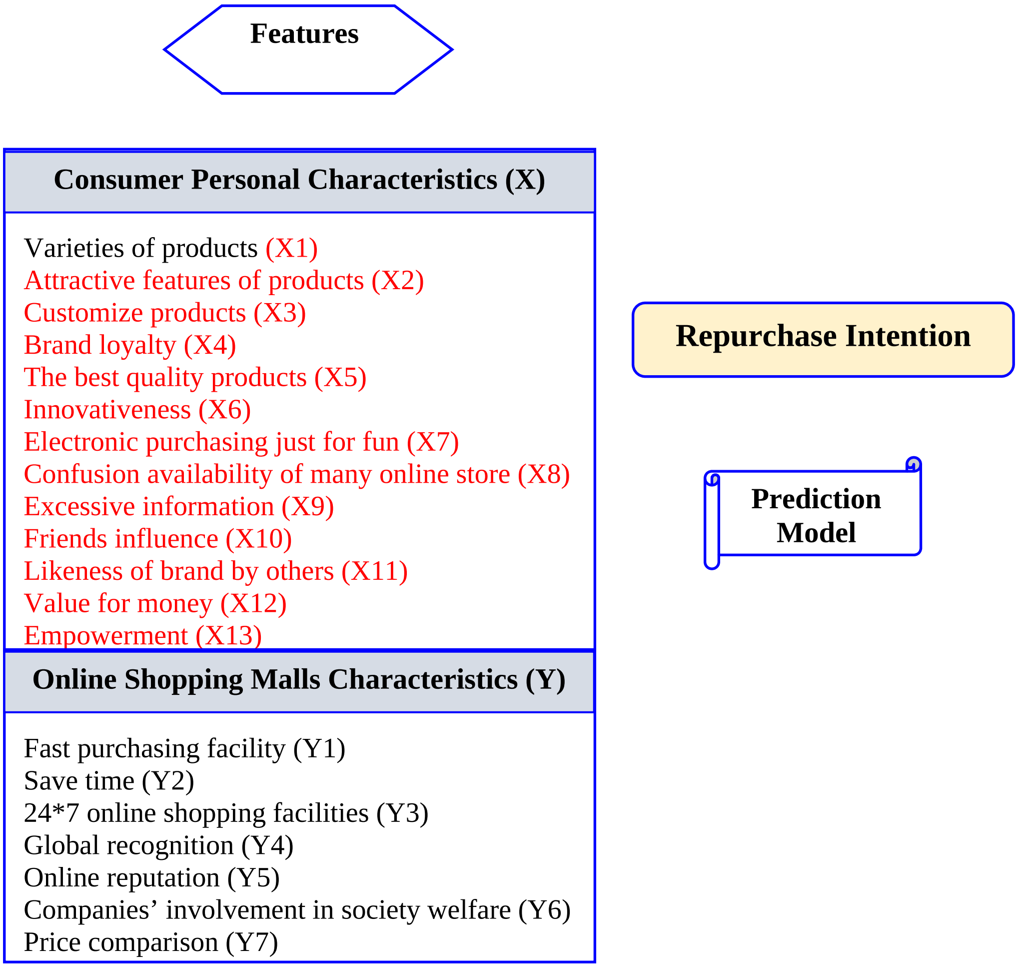


Figure 1: Prediction model

3 Data and methods

This study is divided into three phases. Initially, shopping mall and consumer characteristic's for repurchase intention has been identified through comprehensive literature review. Secondly, ABC has been used to determine the feature selection of consumers' characteristics and shopping malls' attributes (with > 0.1 threshold value). The classification models viz., Decision Trees (C5.0), AdaBoost, RF, SVM and NN, are utilized for prediction of consumer purchase intention. Data has been collected through offline and online methods with a structured questionnaire using 1-5 scale which was sent to the respondents through e-mail and social networking sites. Due to high internet penetration levels and experiences of online purchasing, target respondents of metro cities in India are chosen this study. In the offline mode, judgment sampling is used to collect data. Judgment sampling is based on certain parameters such as, people who have made a purchase online and have internet experience. Pre-testing of the questionnaire has been carried out to examine the validity and applicability of the measure; and final prediction model was developed as shown in Figure 1. Out of 430 questionnaires, 310 questionnaires were collected online and 120 from the offline survey.

Nineteen were eliminated due to inappropriate response. Demographic details of respondents are depicted in Table 2 and online experience is tabulated in Table 3.

Table 2 Demographic Analysis

		Number of Respondents	**Percentage**
Gender	Male	273	66.4
	Female	138	33.6
Age	16-20	52	12.7
	21-25	301	73.2
	26-30	56	13.6
	31-35	2	0.5
Education	Under Graduate	34	8.3
	Graduate	271	65.9
	Post Graduate	99	24.1
	If other specific	7	1.7
Area	Urban	363	88.3
	Rural	48	11.7

Table 3 Internet and Online experiences of the respondents

		Number of Respondents	**Percentage (%)**
Basic purpose of access internet	Study	22	5.4
	Social	48	11.7
	E-mail	72	17.5
	Combination of	269	65.5
A day internet accessibility	Less than 2 times	53	12.9
	2 to 4 times	242	58.9
	More than 5	96	23.4
How long, using the internet?	2 to 3 years	202	49.1
	3 to 4 years	75	18.2
	5 to 6 years	84	20.4
	More than 6	50	12.2
Purchase on the internet?	Less than 1 year	130	31.6
	2 to 3 years	180	43.8
	4 to 5 years	60	14.6
	More than 5	16	3.9

3.1 Artificial Bee Colony (ABC)

Artificial Bee Colony is one of swam knowledge based algorithms and is motivated by the intellectual food scrounging conduct of bumble bees and every result is known as a nourishment hotspot for bumble bees. Condition is firmed taking into account the nature of the food source. Honey bees are classified into onlooker bees, return on bees, and scout bees. Employed bees number is one of the onlooker bees and hunt out the sustenance source and gather data. Onlooker honey bee stays in the hive and explore for the nourishment sources on the premise of data assembled by the employed honey bees. Scout bees exploit new sources of food randomly in abandoned places. For getting a robust solution, ABC takes iteration and work on three phases. Three kinds of parameters can be seen that help in the search process. Process of ABC is given below.

Algorithm 1:
Initialize the parameters;
while Termination criteria is not satisfied **do**
Step 1: Employed honey bee stage for producing new nourishment sources;
Step 2: Onlooker honey bees stage for overhauling the nourishment sources relying upon their nectar content;
Step 3: Scout honey bee stage for finding new sustenance sources set up of deserted nourishment sources;
Step 4: Remember the best sustenance source discovered in this way;
end while
Output the best solution found so far.

3.2 Machine learning methods

Five machine learning classification models, as shown in Table 4, have been used for prediction. Table 5 characterizes the parameter. Small details of each model are given as:

(1) Decision Trees (C5.0): It is an extended version of C45 classification algorithms defined by Quinlan (1986).

(2) AdaBoost: This algorithm is a simple, efficient, and easy-to- use approach to building models (Hastie et al., 2005).

(3) Random Forest (RF): This is grounded on a trees in a forest using random inputs (Liaw and Wiener, 2002).

(4) Support Vector Machine (SVM): It is an influential technique for nonlinear classification and detection of outliers' with an instinctual model representation (Keerthi and Gilbert, 2002).

(5) Neural Network (NN): Neural network training using back-propagation, resilient back-propagation or the modified globally convergent version (Riedmiller and Braun (1993).

Table 4 Machine learning classification model used

Model	**Method**	**Package**	**Parameters**	**References**
Decision Trees	C5.0	C50	winnow, model, trials	Quinlan (1986)
Ada Boost	Ada	logit boost	Iter = 50, bag.frac = 0.5	Hastie et al. (2005)
Random Forest	rf	random forest	mtry	Liaw and Wiener (2002)
SVM	svm	e1071	nu, epsilon	Keerthi and Gilbert (2002)
Neural Network	neuralnet	neuralnet	layer10, layer10, layer10	Riedmiller and Braun (1993)

Table 5 Parameter setting for models

Model	**Parameter Setting**
Decision Trees	Min Split = 20, Max Depth = 30, Min Bucket = 7
Ada Boost	Number of variables = 2, Number of tree = 50
Random Forest	Kernel Radial Basis
SVM	Multinomial
Neural Network	layer10, layer10, layer10

Parameters for measuring classification performance:

(1) Area Under the ROC Curve (AUC): ROC explains the binary classifier system performance when its discrimination threshold is varied. At different threshold settings, the ROC is calculated with true positive rate (TPR) against false positive rate (FPR). AUC normally takes values between 0.5 and 1 (Anagnostopoulos et al., 2012).

(2) Gini: This is a measure of statistical dispersion aimed to represent the target class distribution. The Gini coefficient, given by 2AUC-1, is sometimes preferred (Anagnostopoulos et al., 2012).

(3) H measure: It provides convenient plotting routines that yield insights into the differences and similarities between various metrics. Higher values indicate better performance (Anagnostopoulos et al., 2012).

(4) AUCH: It measures variation of the Area Under the ROCH (Anagnostopoulos et al., 2012).

(5) Kolmogorov-Smirnov Statistic (KS): KS measure the sensitivity and also attains an intuitive graphical interpretation as the maximum vertical distance between the ROC and the diagonal.

4 Methodology

The methodology contains of seven stages as presented in Figure 2. In the first stage, study's aim is defined. The second phase includes features identification through literature review (Section 2) which are related to the consumer as well as shopping malls characteristics. Third phase is all about development of interview questions. In the fourth phase, data is collected through structured questionnaire (described in Section 3). Reliability of the data is performed in phase fifth and the filtering through correlation analysis. In the sixth phase, ABC algorithm is applied. The seventh phase deals about the training and testing of the five machine learning models are trained and tested on the selected data set. Robustness of models has been checked through K-fold in the last stage.

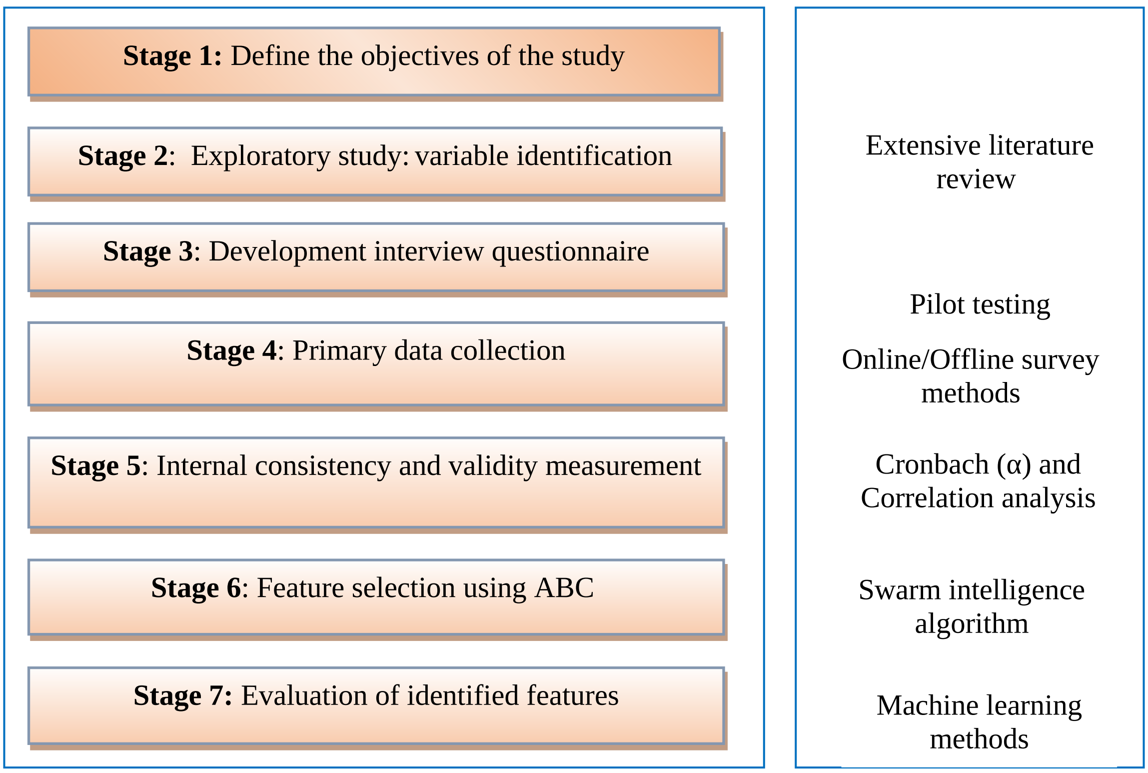


Figure 2: Research Methodology Process

4.1. Internal Consistency and Validity Measurement

Before using ABC for features selection, Eq. (1) has been used to calculate consistency:

$$\alpha = \frac{K}{K-1}\left(1 - \sum_{i-1}^{K} \frac{\delta_Y^2}{\delta_X^2}\right) \tag{1}$$

Where K represents the number of items, δ_X^2 and δ_Y^2 is for measurement of variance. Coefficient of reliability *i.e.*, Cronbach alpha, prefers high ≥ 0.70 (Nunnally, 1978), and to measure convergent validity prefer > 0.50, as mentioned in Table 6. Thus it can be wrapped up that the features in Table 6 have a high consistency, or are fit for use in data collection.

Table 6 Item-Total Statistics

S. No.	Features	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Over all Cronbach's Alpha (α)
1.	X_1	0.554	0.884	0.891
2.	X_2	0.545	0.885	
3.	X_3	0.444	0.888	
4.	X_4	0.351	0.890	
5.	X_5	0.476	0.887	
6.	X_6	0.482	0.887	
7.	X_7	0.509	0.886	
8.	X_8	0.547	0.885	
9.	X_9	0.687	0.880	
10.	X_{10}	0.467	0.887	
11.	X_{11}	0.425	0.888	
12.	X_{12}	0.385	0.889	
13.	X_{13}	0.582	0.884	
14.	X_{14}	0.432	0.888	
15.	X_{15}	0.627	0.882	
16.	X_{16}	0.537	0.885	
17.	X_{17}	0.635	0.882	
18.	X_{18}	0.639	0.882	
19.	X_{19}	0.584	0.887	
20.	X_{20}	0.542	0.891	

4.2 Correlations and Measure of Sample Adequacy

With help of Eq. (2) below, the correlation between the explanatory variables has been collected to measure the relationship within the variable.

$$Correlation = \frac{\sum_{i=1}^{n} (x_i - \acute{x})(y_i - \acute{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \acute{x})^2 \sum_{i=1}^{n} (x_i - \acute{x})^2}} \tag{2}$$

Correlation lies [0, 1] and value towards 1 (Field, 2009) is considered good. The correlation of each feature calculated by Eq. (2) showed that features have significant co-efficient correlation with each other and lies between 0.25 and 0.85, indicating that there is no problem of prefect correction and weak correction among the features (Field, 2009). At this stage we were not able to reduce or skip any feature because each has positive significant co-efficient correlation. Statistical Package SPSS version 21.0 was used to calculate the Anti-image Correlation Matrix to measure sampling adequacy for all features and anti-image is highlighted in the diagonal of Table 7 are above the acceptable level of 0.50 (Field, 2009).

Table 7 Correlations and Measure of Sample Adequacy

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
X_1	**(.860)**																			
X_2	.540**	**(.894)**																		
X_3	.633**	.357**	**(.916)**																	
X_4	.534**	.371**	.378**	**(.862)**																
X_5	.272**	.566**	.304**	.335**	**(.818)**															
X_6	.142**	.347**	.246**	.254**	.394**	**(.785)**														
X_7	.706**	.415**	.497**	.467**	.399**	.349**	**(.861)**													
X_8	.720**	.405**	.626**	.479**	.370**	.395*	.685**	**(.886)**												
X_9	.421**	.374**	.312**	.684**	.322**	.327**	.407**	.384**	**(.841)**											
X_{10}	.213**	.369**	.293**	.254**	.466**	.099*	.308**	.392**	.337**	**(.907)**										
X_{11}	.388**	.383**	.461**	.342**	.336**	.411*	.258**	.294**	.295**	.566**	**(.927)**									
X_{12}	.362**	.323**	.383**	.342**	.445**	.427**	.292**	.450**	.372**	.492**	.497**	**(.860)**								
X_{13}	.407**	.474**	.437**	.452**	.662**	.365**	.453**	.390**	.410**	.377*	.351**	.302**	**(.923)**							
X_{14}	.338**	.342**	.337**	.322**	.390**	.687**	.312**	.335**	.329**	.320*	.352**	.294**	.431**	**(.816)**						
X_{15}	.231**	.236**	.359**	.207**	.213*	.274	.280**	.290**	.264**	.593**	.587**	.235**	.243**	.294*	**(.809)**					
X_{16}	.251**	.420**	.230**	.237**	.376**	.617**	.277**	.184**	.232**	.194*	.219**	.215**	.421**	.589**	.182**	**(.894)**				
X_{17}	.288**	.173**	.411**	.241**	.154**	.073	.293**	.273**	.242**	.595**	.650**	.258**	.138**	.136**	.709**	.186**	**(.862)**			
X_{18}	.241**	.119*	.292**	.221**	.131**	.078	.167**	.212**	.200**	.419**	.521**	.222**	.148**	.099*	.462**	.347**	.508**	**(.925)**		
X_{19}	.262**	.155**	.358**	.301**	.173**	.106*	.218**	.253**	.255**	.539**	.691**	.261**	.140**	.096	.576**	.341**	.639**	.548**	**(.920)**	
X_{20}	.214**	.321*	.376**	.281**	.364**	.321*	.405**	.320**	.279**	.545**	.675**	.302**	.156**	.205*	.605**	.251**	.635**	.562**	.736**	**(.916)**

**. Correlation is significant at the 0.01 level (2-tailed). ; *. Correlation is significant at the 0.05 level (2-tailed).
In diagonal the value shows parenthesis ( ) is Measure of Sample Adequacy (MSA)

4.3 Feature Importance using ABC

ABC assigns optimum weights to every feature owing to the defined objective function, as in Eq. (3). The prominent parameters in ABC are colony size, dimension of the problem, number of food sources, limits and the termination criteria. In this computational endeavor, 2000 number of iterations have been considered.

$$Obj\ fun = min\ ¿ \tag{3}$$

where, T, P, R and n represents the number of instances in training the data set, physical and c properties, the feature and denotes number of properties (eight in this case) respectively. W denotes the weight given to each feature defined in [0, 1]. The weights given to each feature after six different runs are tabulated in Table 8. The average energy weight is highest while area has lowest average value which signifies the each feature importance in the data set. The same can be inferred from Table.8. Since, the weight given to each feature is optimum, all the features were selected for the experiment.

$$\text{Repurchase Intention} \sim X_3 + X_5 + X_6 + X_{12} + Y_2 + Y_3 + Y_5 + Y_7 \tag{4}$$

Where X_3 = Customize products, X_5 =The best quality products, X_6 = Innovativeness, X_{12} = Value for money, Y_2 = Save time, Y_3 = 24*7 online shopping facilities, Y_5 = Online reputation, Y_7 = Price comparison.

After six runs of ABC analysis and taking the average of all six iterations, the features were ranked according their weight (Table 8).Features with < 0.1 threshold value were not given any rank and were not considered for model prediction through machine learning approaches. Brief description of selected features by ABC is given in Table 8.

5 Model evaluation

Performance of the classifiers can be measured by different evaluation metrics depending on specific application. In this study, sensitivity and classification accuracy (S, C) have been considered for evaluating performance of the concerned machine learning classifiers. These metrics are determined by the classification output that comes from confusion matrix. In this matrix, diagonal elements show the object similar to the actual label whereas, off diagonals tells the misclassification information of the model. Let us consider n classes; then confusion C_{ij} of n × n defines the number of pattern of class i predicated in j and Eq. (5) for calculation of accuracy.

$$Accuracy = \frac{\sum_{i=¿¿} n C_{ij}}{\sum_{i=1}^{n} \sum_{i=1}^{n} C_{ij}} \tag{5}$$

Table 8 Importance of each feature using ABC

Runs	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**	**X_7**	**X_8**	**X_9**	**X_{10}**	**X_{11}**	**X_{12}**	**X_{13}**	**Y_1**	**Y_2**	**Y_3**	**Y_4**	**Y_5**	**Y_6**	**Y_7**
1	0.009	0.002	0.149	0.005	0.121	0.011	0.172	0.144	0.011	0.013	0.006	0.156	0.015	0.004	0.002	0.165	0.005	0.142	0.002	0.154
2	0.016	0.011	0.156	0.010	0.144	0.011	0.175	0.142	0.012	0.011	0.010	0.145	0.011	0.001	0.003	0.176	0.009	0.178	0.003	0.143
3	0.021	0.001	0.224	0.011	0.176	0.001	0.181	0.256	0.010	0.011	0.011	0.223	0.011	0.010	0.009	0.247	0.010	0.175	0.004	0.233
4	0.010	0.013	0.223	0.010	0.223	0.005	0.184	0.203	0.012	0.002	0.010	0.256	0.013	0.012	0.010	0.256	0.010	0.254	0.011	0.243
5	0.016	0.011	0.246	0.012	0.255	0.010	0.217	0.214	0.012	0.001	0.012	0.256	0.012	0.010	0.011	0.265	0.012	0.265	0.010	0.242
6	0.050	0.004	0.275	0.001	0.278	0.021	0.228	0.216	0.001	0.003	0.013	0.228	0.011	0.013	0.005	0.275	0.011	0.286	0.011	0.266
Avg.	0.020	0.007	**0.212**	0.008	**0.200**	0.010	**0.193**	**0.196**	0.010	0.007	0.010	**0.211**	0.012	0.008	0.007	**0.231**	0.010	**0.217**	0.007	**0.214**
Rank	-		4	-	6	-	8	7	-	-	-	5	-	-	-	1	-	2	-	3

Table 9 Feature after ABC analysis and their description

Features	**A brief description**
Customize products (X_3)	This factor characterizes customize products seekers, who find seeking out new things pleasurable according their standard and expectations in their products and much Customize conscious about the varieties of products, attractive features, innovative style products and up to date customize products but all according to their needs.
The best quality products (X_5)	This is the characteristic of consumers who are much conscious about quality.
Fast purchasing (X_7)	Consumers want to save time and energy they want quick and time deliver purchasing. Their consciousness is about services.
24*7 online shopping facilities (X_8)	This feature is characterized by a consumer who wants anytime and anywhere purchasing.
Save time (X_{12})	This feature is characterized by a consumer who wants to save their time and also want to enjoy shopping as a fun. Customers believe in save time and try to use it in another productive work.
Online reputation (Y_3)	The degree of consumer consideration about Online reputation of online shopping mall, about their global

	recognition and centralized distributed reputation systems. Perceptions regarding an online store and collectively characterized as an ‘‘online store image.’’
Price comparison (Y_5)	A facility to see different lists of prices for specific products. This feature is related to the price consciousness of the consumer.
Value for money (Y_7)	This is the characteristic of consumers about their consciousness of value for money

Sensitivity is defined by Eq. (6) given by Caballero et al. (2010).

$$S_i = \frac{C_{ij}}{\sum_{j=1}^{n} C_{ij}} \tag{6}$$

Where, n represents classes and c is the matrix and the average can be calculated by Eq. (7) (Caballero et al., 2010).

$$S = \frac{1}{n}\sum_{1=1}^{n} S_i$$

(7)

The accuracy (A) for the classifier is defined in Eq. (8):

$$A = \frac{100}{n}\sum_{1=1}^{n} q_i \tag{8}$$

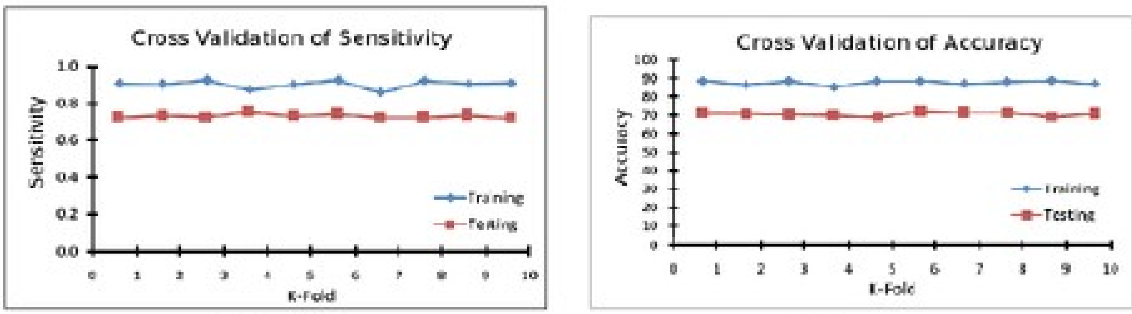
$$q_i = \begin{cases} 1, & \text{if } p_i = a_i \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

Where ‘a’ and ‘p’ are actual and predicted target and n is the total number of instances

Table 10
(70-30%) of data: Models Performances

	Parameters for measurements models								
Models	H	Gini	AUC	AUCH	KS	Accuracy	Sensitivity	ROC	Error
Decision Trees	0.0000	0.0000	0.5000	0.5000	0.0000	93.7742	0.9177	0.5000	0.0343
Ada Boost	**0.2798**	**0.3554**	**0.6777**	**0.8017**	**0.4848**	**97.5806**	**0.9558**	**0.6777**	**0.0242**
Random Forest	0.1649	0.2672	0.6336	0.7355	0.3554	96.7742	0.9477	0.6336	0.0323
SVM	0.1321	0.0854	0.5427	0.7176	0.3636	95.5806	0.9358	0.4573	0.0282
Neural Network	0.0434	0.1818	0.5909	0.5909	0.1818	90.3226	0.8832	0.4091	0.0968

For the measurement of accuracy of predictive models, k-fold cross validation has been utilized and presented Figure 3. Original samples are randomly divided into k sub-samples of the same size. For the testing of the model, first a single sub-sample is retained, and the left k-1sub-samples have been used as training data. For producing a single estimation, average value is calculated and validation on the basis of random sub-sampling.



(a) Sensitivity (b) Accuracy

Figure 3 10-fold presentation

6 Experimental results

The prediction results of five machine learning classification models are analyzed here. All the five classification models were run on their default parameters (see Table 5). Table 10 illustrates the Accuracy (calculated using Eq. (8)) for all the models on 70-30 training-testing partitions. From the results, it is evident that, Ada Boost has the highest sensitivity of (0.95%) and accuracy of (97.58%) on the training-testing partitions respectively. The results validate that Ada Boost outperforms the other learning models for predicting the intention of the consumer in the context of shopping malls' attributes and consumer characteristics. The study explores the machine learning classification models and Artificial Bee Colony (ABC) algorithm, to predict online consumer intention on the basis of consumers' characteristics and online shopping malls' attributes.

Artificial Bee Colony (ABC) algorithm results show that, in consumers' characteristics: (1) Customize products (X_3), (2) The best quality products (X_5), (3) Innovativeness (X_6), (4) Value for money (X_{12}) and in shopping malls' characteristics: (1) Save time (Y_2), (2) 24*7 online shopping facilities (Y_3), (3) Online reputation (Y_5), (4) Price comparison (Y_7) are the important features for consumer during online buying. The rank of these features are: (Y_3) > (Y_5) > (Y_7) > (X_3) > (X_{12}) > (X_5) > (X_6) according to threshold value > 0.01 (the parameters of important features selection through Artificial Bee Colony (ABC)). The model evaluation result indicates that Ada Boost classification outperforms other existing classification models.

7. Conclusion

Theoretical perspective, this study is narrow down the gap from literature in predicting online consumer repurchase intention within the context of shopping malls' and consumers' characteristics using intelligent techniques. The study starts with the problem that how consumers are shifting offline to online and how challenge is increasing for e-service providers to predict their repurchase intention. The study is a revolutionary attempt to combine both characteristics of consumer as well as shopping malls in one model to predict consumer's repurchase intention in the online platform. The prediction model has been developed through extensive literature review and citations. After developing a prediction model, interview questionnaire has been developed and tested through pilot testing. Primary data collection is carried out through online and offline methods. Internal consistency and validity of the data set are measured by Cronbach (α). ABC has been used to find the most important features of shopping malls' and consumers' characteristics and finalized eight characteristics and their ranks. Practically, as for the importance of features selection through ABC, this study practically provides an optimal solution to e-vendors to minimize their targeting cost of consumer as well as to enhance their own services according to the requirements of the consumers. With important features including, customized products, the best quality products, innovativeness, value for money of consumers' characteristics and save time, 24*7 online shopping facilities, online reputation, price comparison of shopping malls' characteristics, the experimental results has been analyzed through five machine learning classification models i.e. Decision Trees (C5.0), AdaBoost, Random Forest (RF), Support Vector Machine (SVM) and Neural Network (NN) in different settings. In the partitions of data into 70-30 training-testing, among all models, the performance of AdaBoost has the highest sensitivity of (0.95%) and accuracy of (97.58%). K-fold cross validation has been used to check sensitivity and accuracy for prediction models. The outcome results of the study may utilize for making marketing strategies more effective and customized according to the requirements of consumers and can generate wealth for the company in the online platform.

References

Akay, B. and Karaboga, D. (2012) ‘A modified artificial bee colony algorithm for real-parameter optimization’, *Information Sciences*, Vol.192, pp.120-142.

Anagnostopoulos, C., Hand, .J., and Adams, N. M. (2012). Measuring classification performance: the hmeasure package. URL: http://cran. r-project.org/web/packages/hmeasure /vignettes/hmeasure. pdf. (Accessed, 18 January, 2016)

Azad, N., Kasehchi, H., Asgari, H., and Bagheri, H. (2014) ‘An exploration study on detecting important factors influencing brand loyalty in retail stores,’ *Decision Science Letters*, Vol.3, No.1, pp.117–120

Azizi, S. and Makkizadeh, V. (2012) ‘Consumer decision-making style: The case of Iranian young consumers,’ *Journal of Management Research*, Vol.4, No. 2, pp.88-102

Blake, B. F., Neuendorf, K. A. and Valdiserri, C. M. (2003) ‘Innovativeness and variety of internet shopping,’ *Internet Research*, Vol.13, No.3, pp.156–169.

Bright, L. F. and Daugherty, T. (2012) ‘Does customization impact advertising effectiveness? an exploratory study of consumer perceptions of advertising in customized online environments,’ *Journal of Marketing Communications*, Vol.18, No.1, pp.19–37.

Cardinale, Y., El Haddad, J., Manouvrier, M. and Rukoz, M. (2010) ‘Web service selection for transactional composition, ‘*Procedia Computer Science*, Vol.1, No.1, pp.2689–2698.

Chan, T. Y., Kadiyali, V., and Park, Y.-H. (2007) ‘Willingness to pay and competition in online auctions,’ *Journal of Marketing Research*, Vol. 44, No.2, pp.324–333.

Chen, Y.-C., Shang, R.-A., and Kao, C.-Y. (2009) ‘The effects of information overload on consumers subjective state towards buying decision in the internet shopping environment,’ *Electronic Commerce Research and Applications*, Vol. 8, No.1, pp.48–58.

Cheong, S. L., and Chin, S. T. S. (2012) ‘Online purchasing behavior among the younger generation,’ *International Journal of Undergraduate Studies*, Vol. 1 No.1, pp.1–7.

Childers, T. L., Carr, C. L., Peck, J., and Carson, S. (2002) ‘Hedonic and utilitarian motivations for online retail shopping behavior,’ *Journal of Retailing*, Vol.77, No.4, pp.511–535.

Citrin, A. V., Sprott, D. E., Silverman, S. N., and Stem Jr, D. E. (2000) ‘Adoption of internet shopping: the role of consumer innovativeness,’ *Industrial management & data systems*, Vol.100, No.7, pp.294–300.

Colantone, I., and Crino`, R. (2014) ‘New imported inputs, new domestic products,’ *Journal of International Economics*, Vol. 92, No.1, pp.147–165.

Cowart, K. O. and Goldsmith, R. E. (2007) ‘The influence of consumer decision-making styles on online apparel consumption by college students,’ *International Journal of Consumer Studies*, Vol. 31, No.6, pp.639–647.

Danaher, P. J., Wilson, I. W. and Davis, R. A. (2003) ‘A comparison of online and offline consumer brand loyalty,’ *Marketing Science*, Vol. 22 No. 4, pp.461–476.

Ellis-Stoll, C. C., and Popkess-Vawter, S. (1998) ‘A concept analysis on the process of empowerment’, *Advances in Nursing Science*, Vol.21, No.2, pp.62–68.

Fernandez Caballero, J. C., Mart´ınez, F. J., Herva´s, C., & Guti´errez, P. A. (2010). Sensitivity versus accuracy in multiclass problems using memetic pareto evolutionary neural networks. *Neural Networks*, IEEE Transactions on, Vol.21, No.5, pp.750–770.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Field, J. M., Heim, G. R., and Sinha, K. K. (2004) ‘Managing quality in the e-service system: Development and application of a process model,’ *Production and Operations Management*, Vol.13, No.4, pp.291–306.

Forbes, L. P., and Vespoli, E. M. (2013) ‘Does social media influence consumer buying behavior? an investigation of recommendations and purchases,’ *Journal of Business & Economics Research*, Vol.11, No.2, pp.107-112.

Gao, J., Zhang, C., Wang, K., and Ba, S. (2012) ‘Understanding online purchase decision making: The effects of unconscious thought, information quality, and information quantity,’ *Decision Support Systems*, Vol.53, No.4, pp.772–781.

Goldsmith, R. E.,and Hofacker, C. F. (1991) ‘Measuring consumer innovativeness,’ *Journal of the Academy of Marketing Science*, Vol.19, No.3, pp.209–221.

Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J. (2005) ‘The elements of statistical learning: data mining, inference and prediction,’ *The Mathematical Intelligencer*, Vol. 27, No.2, pp.83–85.

He, H., Li, Y. and Harris, L. (2012) ‘Social identity perspective on brand loyalty,’ *Journal of Business Research*, Vol. 65, No.5, pp.648–657.

Herbes, C., and Ramme, I. (2014) ‘Online marketing of green electricity in Germany a content analysis of provider’s websites,’ *Energy Policy*, Vol.66, pp.257–266.

Hung, Y.-H., Huang, T.-L., Hsieh, J.-C., Tsuei, H.-J., Cheng, C.-C., and Tzeng, G.-H. (2012) ‘Online reputation management for improving marketing by using a hybrid mcdm model,’ Knowledge-Based Systems, Vol.35, pp.87–93.

Hunt, D. M., Radford, S. K., and Evans, K. R. (2013) ‘Individual differences in consumer value for mass customized products,’ *Journal of Consumer Behaviour*, Vol.12, No.4, pp.327–336.

Iyengar, R., Han, S., and Gupta, S. (2009) ‘Do friends influence purchases in a social network?,’ *Harvard Business School Marketing Unit Working Paper*, pp.09-123.

Jeng, S. P. (2011) ‘The effect of corporate reputations on customer perceptions and cross-buying intentions,’ *The Service Industries Journal*, Vol.31, No.6, pp.851–862.

Jeng, S.-P. (2008) ‘Effects of corporate reputations, relationships and competing suppliers’ marketing programmes on customers’ cross-buying intentions,’ *The Service Industries Journal*, Vol. 28, No.1, pp.15–26.

Jin, C.H. (2013) ‘The effects of individual innovativeness on users adoption of internet content filtering software and attitudes toward children’s internet use. *Computers in Human Behavior*, Vol. 29, No.5, pp.1904–1916.

Johnson, E. J., Moe, W. W., Fader, P. S., Bellman, S., and Lohse, G. L. (2004) ‘On the depth and dynamics of online search behavior,’ *Management Science*, Vol.50, No.3, pp.299–308.

Johnston, K. M. (2001) ‘Why e-business must evolve beyond market orientation: applying human interaction models to computer-mediated corporate communications,’ *Internet Research*, Vol.11, No.3, pp.213–225.

Kamis, A., Stern, T., and Ladik, D. M. (2010) ‘A flow-based model of web site intentions when users customize products in business-to-consumer electronic commerce,’ *Information Systems Frontiers*, Vol.12, No.2, pp.57–168.

Kang, F., and Li, J. (2015) ‘Artificial bee colony algorithm optimized support vector regression for system reliability analysis of slopes,’ *Journal of Computing in Civil Engineering*, pp.04015040.

Kang, F., Li, J., and Li, H. (2013) ‘Artificial bee colony algorithm and pattern search hybridized for global optimization,’ *Applied Soft Computing*, Vol.13 No.4, pp.1781–1791.

Kang, F., Li, J., and Ma, Z. (2011) ‘Rosenbrock artificial bee colony algorithm for accurate global optimization of numerical functions,’ *Information Sciences*, Vol.181, No.16, pp.3508–3531.

Kang, F., Li, J., and Xu, Q. (2009) ‘Structural inverse analysis by hybrid simplex artificial bee colony algorithms,’ *Computers & Structures*, Vol.87, No.13, pp.861–870.

Kang, Y. J., and Lee, W. J. (2015) ‘Self-customization of online service environments by users and its effect on their continuance intention,’ *Service Business*, Vol.9, No.2, pp.321-342.

Karaboga, D., & Akay, B. (2009). A comparative study of artificial bee colony algorithm. Applied mathematics and computation, 214(1), 108-132.

Karaboga, D., & Basturk, B. (2008). On the performance of artificial bee colony (ABC) algorithm. Applied soft computing, 8(1), 687-697.

Karaboga, D., and Akay, B. (2009) ‘A comparative study of artificial bee colony algorithm. *Applied Mathematics and Computation*,’ Vol. 214, No.1, pp.108132.

Keerthi, S., and Gilbert, E. (2002) ‘Convergence of a generalized SMO algorithm for SVM classifier design. *Machine Learning*,’ Vol.46, No.1, pp.351–360.

Khare, A. (2011) ‘Mall shopping behavior of Indian small town consumers,’ *Journal of retailing and consumer services*, Vol.18, No.1, pp.110–118.

Kim, D. J., Ferrin, D. L., and Rao, H. R. (2008) ‘A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents,’ *Decision Support Systems*, Vol.44, No.2, pp.544–564.

Kim, H.W., and Gupta, S. (2009) ‘A comparison of purchase decision calculus between potential and repeat customers of an online store,’ *Decision Support Systems*, Vol.47, No.4, pp.477–487.

Kim, Y.K., Lee, M.-Y., and Park, S.H. (2014) ‘Shopping value orientation: Conceptualization and measurement,’ *Journal of Business Research*, Vol.67, No.1, pp.2884–2890.

Kochukalam, C. A., & Peters, M. J. (2016). Shopping cart experience-Building consumer experience at the check-out stage for online buying. *International Journal of Research in Social Sciences*, *6*(1), 483-490.

Kumar, A., & Dash, M. K. (2015). E-service quality dimensions' effect on customers' willingness to buy: structural equation modelling approach. *International Journal of Services and Operations Management*, *22*(3), 287-303.

Kumar, A., & Dash, M. K. (2016). Using DEMATEL to construct influential network relation map of consumer decision-making in e-marketplace. *International Journal of Business Information Systems*, *21*(1), 48-72.

Kumar, A., and Dash, M. (2013) ‘Constructing a measurement in service quality for Indian banks: Structural equation modeling approach,’ *Journal of Internet Banking and Commerce*, Vol.18, No.1, pp.1-18.

Kumar, A., and Dash, M. (2014) ‘Factor exploration and multi-criteria assessment method (AHP) of multi-generational consumer in electronic commerce,’ *International Journal of Business Excellence*, Vol.7, No.2, pp. 213-236.

Lee, E.-J., and Shin, S. Y. (2014) ‘When do consumers buy online product reviews? Effects of review quality, product type, and reviewers photo,’ *Computers in Human Behavior*, Vol.31, pp.356–366.

Lee, H.-H., and Chang, E. (2011) ‘Consumer attitudes toward online mass customization: an application of extended technology acceptance model,’ *Journal of Computer-Mediated Communication*, Vol.16, No.2, pp.171–200.

Lee, M. Y., Kim, Y.-K., and Fairhurst, A. (2009) ‘Shopping value in online auctions: their antecedents and outcomes, *Journal of Retailing and Consumer Services*, Vol.16, No.1, pp.75–82.

Li, M., Huang, L., Tan, C.-H., and Wei, K. K. (2013) ‘Helpfulness of online product reviews as seen by consumers: Source and content features,’ *International Journal of Electronic Commerce*, Vol.17,No.4, pp.101–136.

Liaw, A., and Wiener, M. (2002) ‘Classification and regression by random Forest,’ *R news*, Vol.2, No.3, pp.18-22.

Lysonski, S., andDurvasula, S. (2013) ‘Consumer decision making styles in retailing: evolution of mindsets and psychological impacts,’ *Journal of Consumer Marketing*, Vol.30, No.1, pp.75–87.

Mahdjoubi, L., Hao Koh, J.,and Moobela, C. (2014) ‘Effects of interactive real-time simulations and humanoid avatars on consumers responses in online house products marketing,’ *Computer- Aided Civil and Infrastructure Engineering*, Vol.29, No.1, pp.31–46.

Maniak, R., Midler, C., Beaume, R.,and Pechmann, F. (2014) ‘Featuring capability: How carmakers organize to deploy innovative features across products,’ *Journal of Product Innovation Management*, VoL.31, No.1, pp.114–127.

Merle, A., Chandon, J.-L., Roux, E., and Alizon, F. (2010) ‘Perceived value of the mass-customized product and mass customization experience for individual consumers,’ *Production and Operations Management*, Vol.19 No.5, pp.503–514.

Merlo, O., Lukas, B. A., and Whitwell, G. J. (2012) ‘Marketing’s reputation and influence in the firm,’ *Journal of Business Research*, Vol.65, No.3, pp.446–452.

Mohanty, R., Ravi, V., and Patra, M. R. (2010) ‘Web-services classification using intelligent techniques,’ *Expert Systems with Applications*, Vol.37, No.7, pp.5484–5490.

Mukherjee, A., Satija, D., Goyal, T. M., Mantrala, M. K., and Zou, S. (2014) ‘Impact of the retail FDI policy on Indian consumers and the way forward,’ In Globalization and Standards (pp. 41-59). Springer India.

Ni, Y., and Fan, Y. (2010) ‘Model transformation and formal verification for semantic web services composition,’ *Advances in Engineering Software*, Vol.41, No.6, pp.879–885.

Nunnally, J. (1978). C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Park, D.-H., and Lee, J. (2009) ‘eWOM overload and its effect on consumer behavioral intention depending on consumer involvement,’ *Electronic Commerce Research and Applications*, Vol.7, No.4, pp.386–398.

Park, Y. A. (2007). *Investigating online decision-making styles* (Doctoral dissertation, Texas A&M University).

Park, Y. A., and Gretzel, U. (2010) ‘Influence of Consumers' online Decision-Making Style on Comparison Shopping Proneness and Perceived Usefulness Of Comparison Shopping Tools,’ *Journal of Electronic Commerce Research*, Vol.11, No.4, pp.342-354.

Pires, G. D., Stanton, J., and Rita, P. (2006) ‘The internet, consumer empowerment and marketing strategies,’ *European Journal of Marketing*, Vol.40, No.9/10, pp.936–949.

Qin, A., Huang, V., and Suganthan, P. (2009) ‘Differential evolution algorithm with strategy adaptation for global numerical optimization, ‘*Evolutionary Computation*, IEEE Transactions on, Vol.13, No.2, pp.398–417.

Quinlan, J. (1986) ‘Induction of decision trees,’ *Machine Learning*, Vol.1, No.1, pp.81–106.

Rajeev, P. V., & Rekha, V. S. (2016). Opinion Mining of User Reviews Using Machine Learning Techniques and Ranking of Products Based on Features. In Proceedings of the International Conference on Soft Computing Systems (pp. 627-637). Springer India.

Riedmiller, M., and Braun, H. (1993) ‘A direct adaptive method for faster backpropagation learning: The RPROP algorithm,’ In *Neural Networks*, 1993, IEEE International Conference on (pp. 586–591).

Rieger, M. O. (2012) ‘Why do investors buy bad financial products? probability misestimation and preferences in financial investment decision,’ *Journal of Behavioral Finance*, Vol.13, No.2, pp.108–118.

Riquelme, I. P., and Roma´n, S. (2014) 'The influence of consumers cognitive and psychographic traits on perceived deception: A comparison between online and offline retailing contexts,' *Journal of Business Ethics*, Vol.119, No.3, pp.405–422.

Rizwan, M., Sultan, H., Parveen, S., Nawaz, S., Sattar, S., and Sana, M. (2013) 'Determinants of online shopping and moderating role of innovativeness and perceived risk,' *Asian Journal of Empirical Research*, Vol.3, No.2, pp.142–159.

Saraswat, M., and Arya, K. (2013) 'Supervised leukocyte segmentation in tissue images using multi- objective optimization technique,' *Engineering Applications of Artificial Intelligence*, Vol.31, pp.44-52.

Seng, J. L., and Chen, T. C. (2010) 'An analytic approach to select data mining for business decision,' *Expert Systems with Applications*, Vol.37, No.12, pp.8042-8057.

Sicilia, M., and Ruiz, S. (2010) 'The effects of the amount of information on cognitive responses in online purchasing tasks,' *Electronic Commerce Research and Applications*, Vol.9, No.2, pp.183–191.

Thirumalai, S., and Sinha, K. K. (2011) 'Customization of the online purchase process in elec- tronic retailing and customer satisfaction: An online field study,' *Journal of Operations Management*, Vol.29, No.5, pp.477–487.

Van der Heijden, H., and Verhagen, T. (2004) 'Online store image: conceptual foundations and empirical measurement,' *Information & Management*, Vol.41, No.5, pp.609–617.

Verhagen, T., and Van Dolen, W. (2009) 'Online purchase intentions: A multi-channel store image perspective,' *Information & Management*, Vol.46, No.2, pp.77–82.

Vyas, R., and Sisodia, R. S. (2013) 'A study of factors influencing consumers' preference of retail stores for fresh consumables in Jaipur,' *EXCEL International Journal of Multidisciplinary Management Studies*, Vol.3, No.6, pp.1–9.

Wang, Y., and Tseng, M. M. (2013) 'Customized products recommendation based on probabilistic relevance model,' *Journal of Intelligent Manufacturing*, Vol.24, No.5, pp.951–960.

Webster Jr, F. E. (1975) 'Determining the characteristics of the socially conscious consumer,' *Journal of Consumer Research*, Vol.2, No.3, pp.188-196.

Wu, C. H., Kao, S.C., and Lin, H. H. (2013) 'Acceptance of enterprise blog for service industry,' *Internet Research*, Vol. 23, No.3, pp.260–297.

Yang, M. H., Weng, S.S., and Hsiao, P.-I. (2014) 'Measuring blog service innovation in social media services,' *Internet Research*, Vol.24, No.1, pp.110–128.

Zhou, L., Dai, L., and Zhang, D. (2007) 'Online shopping acceptance model-a critical survey of consumer factors in online shopping,' *Journal of Electronic Commerce Research*, Vol.8, No.1, pp.41-52

Zhou, Q., Xia, R., & Zhang, C. (2016). Online shopping behavior study based on multi-granularity opinion mining: China versus America. Cognitive Computation, 1-16.

Zinkhan, G. M., and Watson, R. T. (1998) 'Electronic commerce: a marriage of management information systems and marketing,' *Journal of Market-Focused Management*, Vol.3, No.1, pp.5–22.